Planetary Ionospheres

Nanan Balan

Department of Automatic Control and Systems Engineering

University of Sheffield, Sheffield S1 3JD, UK

Abstract: The paper presents a summary of the lectures on *planetary ionospheres* given at NASA's 1st Asia Pacific School on International Heliophysical Year conducted at Indian Institute of Astrophysics, Kodaikanal, India during 10-22 December 2007. Following an introduction, the paper describes the structure of the ionospheres, theory of Earth's ionosphere including the effects of diffusion, neutral wind and electric field, and ionospheric electric fields, currents and variability.

1. Introduction

Ionosphere is the ionized part of the upper atmosphere. In the case of Earth, ionosphere extends from about 70 to 1000 km height, with peak at around 300 km on the average. It is a weakly ionized plasma with ionized to neutral particle density ratio reaching a maximum of about 1:500. The region above ionosphere is plasmasphere, above which is magnetosphere. The ionosphere is embedded in neutral atmosphere (mesosphere-thermosphere region), the density of which decreases exponentially with height. There are several books on ionosphere (e.g., Rishbeth and Garriott, 1969; Ratcliffe, 1972; Mahajan and Kar, 1988; Kelley, 1989).

The ionosphere has been studied for its scientific and practical importance. It influences the propagation of electromagnetic waves and hence has advantages and disadvantages in communication. It is to exploit the advantages in radio communication and find remedies for the disadvantages in all branches of tele-communication that the ionosphere has been studied since Marconi's successful transmission of radio signals across the Atlantic. Since ionosphere varies from place to place and from time to time, it is important to study it at as many places as possible and for as long as possible. In this context, ionospheric variations during geomagnetic storms form an integral part of *Space Weather*. Ionospheric studies are also important for radio astronomy, to account for the irregular changes in the amplitude and phase (or scintillations) of the radio signals that pass through the ionosphere.

On the scientific side, much of the basic science has been understood. Nevertheless, the ionosphere is being studied with renewed interest using radio and in-situ techniques to understand the science of many interesting phenomena that occur in it and also to understand its interactions with the regions above and below. In addition, the ionosphere is a naturally existing plasma. We can study it at our will and wish, and use the information to understand the problems associated with man-made plasmas, in fusion reactors. The

existence of the ionosphere also results in the generation of natural resonators like the *ionospheric waveguide* formed around the plasma density maximum for the propagation of magnetosonic waves, *Alfven resonator* formed between the density maximum and an upper altitude at about 3000 km for Alfven waves, and *Schumann resonator* formed between the nearly perfectly conducting terrestrial surface and ionosphere.

2. Ionospheric structure

All planets (except Mercury) have ionospheres. The density of Earth's ionosphere increases with height, reaches a maximum (ionospheric peak) and then decreases for further increase in height (Figure 1). The vertical structure also includes regions of enhanced density, called D, E and F regions during daytime. The D region occurs below 90 km, E region between 90 and 150 km and F region at heights above 150 km. At night, the ionosphere reduces to a single layer structure called F layer. A weak C region sometimes forms near 70 km. Comparatively strong ionization, populated by ion hydrates, sometimes occurring below 70 km is called cosmic ray layer. The regions below and above the ionospheric peak are referred to as bottomside and topside ionospheres. The electron densities in the E and F regions are high enough to reflect medium and short wavelength radio waves used for long distance radio communications. The densities in the D region and below are not sufficient to reflect radio waves for communication; instead, these regions absorb radio waves.

Figure 2 shows the electron density profiles of Venus's ionosphere at solar maximum and minimum (Knudsen et a., 1987). As shown, the ionosphere is nearly equally strong at solar maximum and minimum, with peak at around 150 km height, though topside ionosphere depletes significantly at solar minimum. The structures of the ionospheres of other planets can be found in Hinson (http://www.star.Stanford.edu/projects/), Atreya et al. (1984) and Mahajan and Kar (1988). Comparing all planets, the ionosphere is stronger in Earth

3

than in other planets.

3. Conservation laws

The behaviour of the ionosphere is controlled by chemical, dynamical and energy processes described by the conservation laws of mass, momentum and energy. The chemical processes determine the production and loss of ionization (ions and electrons) while dynamical processes determine movement of ionization. The rate of change of ionization in the ionosphere can be obtained from the *continuity equation* (or law of conservation of mass)

$$\frac{\delta n_e}{\delta t} = Q - L - \bigtriangledown.(n_e \mathbf{V}) \tag{1}$$

where Q, L and n_e are the rate of production, rate of chemical loss and number density of electrons; $\bigtriangledown.(n_e \mathbf{V})$ is the loss of ionization due to transport, $\mathbf{V}$ being transport velocity.

The motion of charged particles in the ionosphere is controlled mainly by the forces of pressure gradient, gravity, electric field and collisions. The *momentum equation*, which balances these forces, for the jth ion species can be written as

$$\rho_j \frac{d\mathbf{V}_j}{dt} = -\bigtriangledown p_j + \rho_j \mathbf{g} + \frac{q_j \rho_j}{m_j}(\mathbf{E} + \mathbf{V}_j \times \mathbf{B}) - \sum_k \rho_j \nu_{jk}(\mathbf{V}_j - \mathbf{V}_k) \tag{2}$$

where $\mathbf{V}$, ρ, q and m are the velocity, density, charge and mass of ion j, ν_{jk} is the collision frequency of ions j and k, $\mathbf{E}$ and $\mathbf{B}$ are electric field and magnetic field, and $j \neq k$. The first term in equ. 2 represents the force due to pressure gradient ($\bigtriangledown p_j$). The terms involving g and $\nu_{jk}(\mathbf{V}_j - \mathbf{V}_k)$ are the forces due to gravity and collision (or friction) between ions j and k, and terms involving $\mathbf{E}$ and $\mathbf{B}$ are the forces due static electric field ($\mathbf{E}$) and Lorentz electric field ($\mathbf{V}_j \times \mathbf{B}$). Since there are different ion species, a complete picture of the plasma motion would require treatment of a large number of coupled equations. The forces of viscosity, Coriolis acceleration, centripetal acceleration, and tides and waves are not included in equation 2.

Theoretical values of electron and ion temperatures can be obtained by solving the *energy equations*. The electron energy equation is

$$\frac{3}{2}n_e k \frac{\delta T_e}{\delta t} = -n_e k T_e \bigtriangledown .V_e - \frac{3}{2}n_e k V_e . \bigtriangledown T_e - \bigtriangledown . \mathbf{q}_e + \sum Q_e - \sum L_e \tag{3}$$

where $\mathbf{q}_e = -\lambda_e \bigtriangledown T_e$ is the electron thermal heat flux, with λ_e being the thermal conductivity coefficient. The first term on the right-hand side represents adiabatic expansion, second term accounts for advection, third term is the divergence of electron heat flow, and $\sum Q_e$ and $\sum L_e$ are the sum of all the heating and cooling rates.

The ion energy equation, similar to the electron energy equation, is given by

$$\frac{3}{2}N_i k \frac{\delta T_i}{\delta t} = -N_i k T_i \bigtriangledown .V_i - \frac{3}{2}N_i k V_i . \bigtriangledown T_i - \bigtriangledown . \mathbf{q}_i + \sum Q_i - \sum L_i \tag{4}$$

where $\mathbf{q}_i$ is ion thermal heat flux, $\sum Q_i$ is the sum of all heating rates (mainly thermal ion heating due to Coulomb interactions with thermal electrons), and $\sum L_i$ is the sum of all the ion cooling rates (mainly to neutral gas).

4. Theory of photoionization

The ionosphere is formed mainly from the photoionization of the atmosphere by solar X-rays and EUV radiations; high energy charged particles from the Sun also contribute to the ionization at high latitudes, which is not considered here. The intensity of solar radiations (X-rays and EUV) decreases due to absorption as they penetrate through the atmosphere while the concentration of the ionizable constituents increases. That causes the ionization to maximize at a height where the intensity of the radiation and concentration of the constituents optimise.

4.1. Simplified production function

Considering a plane horizontally stratified atmosphere composed of a single gas constituent on which a monochromatic solar radiation is incident at a zenith angle (angle with

the vertical) χ, the rate of production of ionization Q is given by (Ratcliffe, 1972)

$$Q = Q_m exp \left[1 + \frac{h_m - h}{H} - exp \left(\frac{h_m - h}{H} \right) \right] \tag{5}$$

where Q_m is the peak rate of production at height h_m and $H = \frac{kT}{mg}$ is the *scale height* of the atmospheric constituent; scale height of a gas is the height at which its density decreases to $1/e$ times its density at a given base level. Using reduced height $z = \frac{h - h_m}{H}$ measured above the peak of production, Q can be written as

$$Q = Q_m exp \left[1 - z - exp \left(-z \right) \right] \tag{6}$$

The simplicity of the expression shows that the production layers corresponding to all wavelengths and solar zenith angles (χ) have same shape in terms of the normalized quantities z and Q/Q_m as shown in Figure 3.

For vertical incidence of solar radiation $(\chi = 0)$, Q becomes

$$Q = Q_0 \, exp \left[1 - y - \sec \chi exp \left(-y \right) \right] \tag{7}$$

where $y = (z - z_0)/H$ with z_0 being the height of peak production Q_0. For very large y $(y >> 0 \ \text{ or } \ z >> z_0)$, Q takes the form $Q = Q_0 \, exp \left(-\frac{z}{H} \right)$. That is, at heights well above the ionization maximum, the ionization decays at the rate determined by the scale height of the ionizable constituent. For very small y $(y << 0 \ \text{ or } \ z << z_0)$, Q becomes $Q = Q_0 \, exp \left[- \sec \chi exp \left(-\frac{z - z_0}{H} \right) \right]$ which decreases rapidly toward lower altitudes.

4.2. Gracing incidence and generalized production function

The plane-earth approximation made above is not valid for χ near $90°$ because $\cos \chi$ reaches zero and therefore there should not be any production (equ. 7) while, in reality, significant production occurs even at $\chi = 100°$ when H is large. To overcome this difficulty, Chapman defined a *grazing incidence function* $Ch(x, \chi)$ as

$$Ch(x, \chi) = \left(\frac{1}{2} \pi x \sin \chi \right)^{1/2} exp \left(\frac{1}{2} \cos^2 \chi \right) \left[1 \pm erf \left(\frac{1}{2} x \cos^2 \chi \right)^{1/2} \right] \tag{8}$$

6

where $x = (R_e + h)/H$ and the error function erf is readily available. The function $Ch(x, \chi)$ should replace $\sec \chi$ near sunrise and sunset. It departs from $\sec \chi$ when $\chi > 80°$. For example, at $90°$, $\sec \chi$ becomes infinite while $Ch(x, 90°)$ becomes $(\frac{1}{2}\pi x)^{1/2}$.

The other assumptions, such as monochromatic solar radiation and single constituent atmosphere, are also not valid. The atmosphere is composed of different gases of different scale heights and ionization potentials, and solar radiation spectrum consists of a myriad of lines and bands with different intensities. Further, the absorption cross section varies with gas species and photon wavelength. The approach is therefore to assume a neutral atmosphere model with height distributions of some of the major gases in concern, and then for a finite number of wavelength bands derive the production profiles for each gas, as used in mathematical models of the ionosphere (e.g., Schunk and Sojka, 1996; Bailey and Balan, 1996). A generalized production function Q_i can be defined as

$$Q_i = \sum_\lambda I(\lambda)\sigma_i(\lambda)n_i \; exp\left(-\sum_j \sigma_j(\lambda)n_j H_j Ch_j(\chi)\right) \tag{9}$$

where $I(\lambda)$ is the intensity of solar radiation of wavelength λ, $\sigma_i(\lambda)$ and $\sigma_j(\lambda)$ are the photoionization and photoabsorption cross-sections of the ith and jth neutral gases of densities n_i and n_j; H_j is the scale height of the jth neutral gas and $Ch_j(\chi)$ is the Chapman function of the jth gas. The summation over λ is for the wavelength range of the ionizing radiation and summation over j is for the important absorbing gases O, O_2 and N_2.

4.3. Optical depth and ionization potential

The intensity of solar radiation decreases due to absorption as it passes through the atmosphere. At the peak of ionization production, the intensity reduces to $1/e$ times the intensity at the top of the atmosphere ($I_m = I_\infty/e$); at that altitude the intensity is said to reach one *optical depth*. The optical depth τ at any altitude where the density is n and scale height is H is given by $\tau = \sigma n H \sec \chi$. Since $\sec \chi$ is least when χ is zero, unit optical

7

depth is reached at the lowest altitude for an overhead Sun. The altitude of unit optical

depth for an over head sun is illustrated in Figure 4 for the solar spectrum below 3000Å.

As shown, the different wavelengths have different heights of unit optical depth.

Production of ionization depends also on the ionization potentials V_p of the different

species. The ionization and dissociation potentials (and equivalent wavelengths, according

to $V_p = h\nu = h\frac{c}{\lambda}$, with h being Plank's constant and ν and λ being photon frequency

and wavelength) of the most abundant species in the atmosphere are listed in Table 1.

The species get ionized or dissociated by a photon if the photon energy ($h\nu$) is greater

than V_p. Typical ionization cross sections in the EUV regime are of the order of 10^{-21} -

10^{-22} m^2. For the low end of this range and an overhead Sun, unit optical depth of EUV

radiation corresponds to $\frac{1}{\sigma} = nH = 10^{22}$ m^{-2}, which is satisfied at about 110 km height.

That explains why most of the solar EUV radiation is absorbed above 100 km altitude.

Photodissociation has a smaller cross-section which satisfies down to 80 km height where

$nH \approx 10^{24}$ m^{-2}. The collisional cross-section for O$_2$ is reduced to 10^{-27} m^2 in the Herzberg

continuum (2026-2424Å). This is the reason why these wavelengths can penetrate deeply

into the stratosphere where $nH = 10^{27}$ m^{-2}, where O$_2$ can dissociate to form O atoms.

4.4. Chemical loss of ionization

Once ionization is produced, a part of it is lost by chemical reactions and another part

is removed (or added) by transport mechanisms. Important loss reactions are radiative and

dissociative recombinations

$$X^+ + e \rightarrow X + h\nu \tag{10}$$

$$XY^+ + e \rightarrow X + Y \tag{11}$$

The radiative recombination (equ. 10) removes O$^+$ ions through $O^+ + e \rightarrow O + h\nu$, with

the liberated photon (needed to conserve energy) having optical wavelength. It is a com-

130 paratively slow reaction because it is difficult to conserve momentum as it results in only

131 one material particle. A more rapid loss process for O^+ is through a chain of reactions

132 involving N_2 or O_2. The dissociative recombination (equ. 11) is important for the loss of

133 N_2^+, O_2^+, and NO^+ ions (primary ions in D and E regions).

At heights where both reactions (10 and 11) are important, charge neutrality requires $n_e = [X^+] + [XY^+]$, and the loss rate L_i should be proportional to the product of electron and ion densities as $L_i = \alpha n_e^2$ where α is the recombination coefficient. Considering photochemical equilibrium (when production balances chemical loss), the continuity equation becomes $Q_i = L_i = \alpha n_e^2 = Q_0 exp\left[1 - y - \sec \chi exp(-y)\right]$. That gives

$$n_e(z) = \sqrt{\frac{Q_0}{\alpha}} exp \left[\frac{1}{2}\left(1 - y - \sec \chi exp(-y)\right)\right] \tag{12}$$

134 This is referred to as α-Chapman profile, which is representative of E region where molecular

135 ions are dominant and transport processes are unimportant.

However, when charge neutrality is maintained with the loss of one electron-ion pair, $L_i = \beta n_e$ where β is the loss coefficient. Photochemical equilibrium then becomes $Q_i = L_i = \beta n_e$, which gives the electron density profile called β-Chapman profile as

$$n_e(z) = \frac{Q_0}{\beta} exp \left[1 - y - \sec \chi exp(-y)\right] \tag{13}$$

136 ## 4.5. Formation of ionospheric regions

137 The ionizing radiations have a wide, continuous spectrum superposed by numerous

138 intense emission lines. The atmospheric constituent have different ionization potentials

139 and, in the thermosphere, they are distributed according to barometric law. Therefore,

140 ionization of the different constituents maximizes at different altitudes. That accounts in

141 part to the formation of the different ionospheric regions (Figure 1).

142 D region is an ion-neutral collision-dominated region of molecular ions (O_2^+, N_2^+, NO^+)

9

produced by X-rays (< 8Å), EUV (1027Å -1118Å) and H Ly-α (1216Å) radiations. Galactic

cosmic rays, the most energetic radiation, provides additional ionization for D region and

below. The cosmic rays are so energetic that the secondary energetic particles associated

with primary cosmic rays penetrate to ground level before reaching the ends of their paths.

The primary and secondary cosmic ray ionization results in the cosmic ray layer (Figure 1).

The E region is produced mainly from the photoionization of O and O_2 by EUV radiation

(796-1027Å), X-rays (8-140Å) and H Ly β (1026Å) line. The chemical loss-process in the

E-region is one of dissociative recombination (discussed below).

The F region, main ionospheric region, contains bulk of the ionosphere. It is produced

mainly from the photoionization of O and partly form N_2 and O_2 by EUV (170-796Å)

radiation. The important chemical loss processes in the F region are those leading to

the loss of O^+ ions through radiative and dissociative recombinations (discussed below).

Although F region ionization production maximizes below 200 km height, the region splits

into F_1 and F_2 layers during daytime due to the altitude variations of chemical loss processes

and dynamical processes (discussed below), F_1 layer near the peak of ionization production

and F_2 layer (main ionospheric layer) at higher heights.

4.6. Formation of F_1 layer

The main ion (O^+) in the F region is lost through the ion-neutral rearrangement reactions $O^+ + N_2 \rightarrow NO^+ + N$ and $O^+ + O_2 \rightarrow O_2^+ + O$. The resulting molecular ions (NO^+ or O_2^+) are lost by comparatively rapid dissociative recombination process with electrons. The height at which the ion-neutral rearrangement and dissociative recombination becomes equal is called transition height h_t, below which rearrangement dominates and above which dissociation dominates. A clear F_1 layer is formed when the transition height h_t lies above h_0 (height of peak ionization production), which is satisfied during daytime away from noon

in the summer of low solar activity. During other daytime periods h_t lies below h_0, and hence no clear F_1 peak is formed at h_0, instead an F_1 ledge is formed (Ratcliffe, 1972).

4.7. Plasma diffusion and F_2 layer

Plasma diffusion can explain the formation of ionospheric peak (F_2 layer) at heights above the peak of ionization production (e.g., Yonezawa, 1956; Ratcliffe (1956). For simplicity, consider vertical plasma diffusion in the absence of magnetic field ($\mathbf{B} = 0$), neutral wind ($\mathbf{U} = 0$) and electric current. Though electrons are lighter, they diffuse together as a single system (ambipolar plasma diffusion) due to the polarization electric field between electrons and ions. The solution of the equations for the vertical motion of electrons and ions, gives the vertical plasma flux (nW) as

$$nW = -\frac{k\left(T_i + T_e\right)}{m_i \nu_i}\left[\frac{dn}{dz} + \frac{nm_ig}{k\left(T_i + T_e\right)}\right] \tag{14}$$

$$= -D_p\left(\frac{dn}{dz} + \frac{n}{H_p}\right) \tag{15}$$

where $D_p = \frac{k(T_i+T_e)}{m_i \nu_i}$ is ambipolar diffusion coefficient and $H_p = \frac{k(T_i+T_e)}{m_i g}$ is plasma scale height. If $T_e = T_i$, then $D_p = 2D$ and $H_p = 2H$ where D and H are the diffusion coefficient and scale height of a gas of neutral particles having the same mass and temperature as ions. The solution of plasma diffusion equation

$$\frac{\delta n}{\delta t} = -\frac{\delta}{\delta z}(nW) = \frac{\delta}{\delta z}\left[D_p\left(\frac{\delta n}{\delta z} + \frac{n}{H_p}\right)\right] \tag{16}$$

under equilibrium becomes $\frac{\delta n}{\delta z} = -\frac{n}{H_p}$, which for constant H_p gives

$$n = n_0 \, exp\left(-\frac{z - z_0}{H_p}\right) \tag{17}$$

where n_0 is the plasma density at a reference altitude z_0, for example, at F region maximum. The equilibrium electron density profile above the ionospheric peak therefore decays exponentially with altitude with a scale height twice as large as the scale height of the neutral

11

atmosphere (because $H_p = 2H$). The diffusion coefficient D_p increases exponentially with increasing altitude as

$$D_p = D_0 \, exp\left(\frac{z - z_0}{H}\right) \qquad (18)$$

170   where D_0 is the diffusion coefficient at z_0.

171   In the ionosphere; T_e is always higher than T_i (except during frictional heating). Let

172   $T_e = c \, T_i$, with $c > 1$. Then D_p and H_p increase by the factor $(1 + c)/2$ when $T_e > T_i$.

173   Thus, the shape of the electron density profile and plasma diffusion are strongly dependent

174   on the electron-ion temperature ratio.

175   **Effect of geomagnetic field**

Plasma flux along geomagnetic field (due to diffusion, from equ. 15) and its vertical component are

$$nv_\parallel = -D_p \sin I \left(\frac{dn}{dz} + \frac{n}{H_p}\right) \qquad (19)$$

$$nW = nv_\parallel \sin I = -\sin^2 I D_p \left(\frac{dn}{dz} + \frac{n}{H_p}\right) \qquad (20)$$

176   The vertical flux becomes zero at $I = 0$ (equator) where field lines are horizontal.

177   **F$_2$ layer**

As discussed in the formation of F$_1$ layer (section 4.6), the electron density above the F$_1$ layer increases (due to slow decrease of the production rate Q compared to chemical loss coefficient β). At high altitudes, where both Q and βn_e are unimportant and electron density is controlled mainly by diffusion, the electron density decreases exponentially with increasing altitude as discussed above. Thus, the electron density must maximize at an altitude in between these altitude regions. That is where the ionospheric peak (or the F$_2$ peak) occurs. Under steady state at low altitudes, $Q = \beta n_e$ (which is β Chapman relation).

At high altitudes, where $Q \approx 0$ and $\beta \approx 0$, the steady state solution gives

$$n = n_{max} \left(-\frac{z - z_{max}}{H_p} \right) \tag{21}$$

$$z_{max} = \frac{H}{1 + \lambda} ln \frac{H H_p \beta_0}{D_0 \sin^2 I} \tag{22}$$

where H is the scale height of a mixture of N_2 and O_2 and $\lambda = m(O_2, N_2)/m(O)$ is their total mass compared to atomic oxygen mass.

4.8. Effect of neutral wind and electric field

As shown above, ionospheric peak occurs at altitudes above the peak of ionization production due to plasma diffusion. The peak can be shifted by neutral air winds and electric fields. A neutral wind of meridional component U_θ and zonal component U_ϕ produces a wind U in the magnetic meridian $U = U_\theta \cos \alpha \pm U_\phi \sin \alpha$ where α is magnetic declination angle and opposite signs correspond to opposite hemispheres. A horizontal wind blowing towards magnetic equator with speed U drives F region plasma up the geomagnetic field lines with a speed $U \cos I$, of which vertical component is $U \sin I \cos I$, with I being dip angle. This upward drift raises the ionospheric peak to altitudes of reduced chemical loss and so increases peak electron density. An opposite effect happens when the wind blows poleward. Very roughly, the rise in peak height $\triangle z$ for an equatorward wind U is

$$\triangle z \approx \frac{H}{D} U \sin I \cos I \tag{23}$$

The altitude of the ionosphere peak can also be drifted by electric field $\mathbf{E}$ at places where magnetic field is not vertical. The drift velocity is $\mathbf{E} \times \mathbf{B} / B^2$, and its vertical component is upward for eastward $\mathbf{E}$ and downward for westward $\mathbf{E}$. This electromagnetic drift is most effective at equatorial latitudes, where the magnetic field is horizontal and $\mathbf{E} \times \mathbf{B}$ drift is totally vertical. The $\mathbf{E} \times \mathbf{B}$ drift is also effective at high latitudes where magnetic field is vertical and hence plasma drifts horizontally.

F$_3$ layer

An additional stratification, called F$_3$ layer, can occur in the equatorial F region during daytime (Balan et al., 1998). As daytime F$_2$ layer drifts upward mainly due to upward $\mathbf{E}\times\mathbf{B}$ drift and partly due to equatorward neutral wind (and other dynamical sources such as tides and waves), a new layer develops at lower altitudes through the usual photochemical and dynamical processes of the equatorial F region. As time progresses, the original F$_2$ layer drifts upward and forms F$_3$ layer while the new layer develops into usual F$_2$ layer. Both layers can be detected by bottomside ionosondes for sometime when the density of F$_3$ layer remains greater than that of F$_2$ layer. As time progresses further, both layers drift upward, and when the density of F$_3$ layer decreases below that of F$_2$ layer, the F$_3$ layer can be observed as topside ledges. Strong F$_3$ layer develops and rapidly drifts to the topside during strong daytime eastward electric field as during prompt penetration of electric field (Figure 5).

5. Ionospheric electric fields and currents

When ions and electrons respond differently to the forces acting on them, there arises an electric current; this current diverges if there exists a spatial variation in electrical conductivity. Then a (polarisation) electric field builds up quickly (within microseconds in the ionosphere) to modify the ion and electron velocities so that current divergence becomes zero. These physical process together with an ionospheric dynamo can explain the existence of electric fields and currents in the ionosphere.

5.1. Dynamo theory

When air (due to tides, waves and winds) tries to move ionospheric plasma across geomagnetic field, there arises an electric field $\mathbf{U} \times \mathbf{B}$ (Lorentz field), which drives currents at levels in the ionosphere (mainly E region) where electrical conductivity is appreciable.

14

The current density $\mathbf{j}$ is given by $\mathbf{j} = \sigma.(\mathbf{U} \times \mathbf{B})$ where σ is conductivity. However, due to the vertical and horizontal variations of σ, currents cannot flow freely in all directions. That causes accumulation of electric charge, and a polarization electric field $\mathbf{E} = -\bigtriangledown \phi$ builds up, ϕ being electric potential. The total electric field $\mathbf{E'} = \mathbf{E} + \mathbf{U} \times \mathbf{B}$. The polarization field adjusts itself until the current flow becomes non-divergent and horizontal. The current density $\mathbf{j'}$ due to the total field is $\mathbf{j'} = \sigma'.\mathbf{E'} = \sigma'.(\mathbf{E} + \mathbf{U} \times \mathbf{B})$. As described below, the dynamo action is effective at E region heights because at these heights electrons and ions respond differently to neutral wind. At F region altitudes, the plasma motion is primarily along geomagnetic field lines.

Ion and electron velocities

The ion and electron velocities (from momentum equation 2) are

$$\mathbf{V}_i = \mathbf{U} + \frac{1}{1 + k_i^2} \left[\frac{k_i}{B}\mathbf{E'} + \left(\frac{k_i}{B}\right)^2 \mathbf{E'} \times \mathbf{B} + \left(\frac{k_i}{B}\right)^3 (\mathbf{E'}.\mathbf{B})\mathbf{B} \right] \tag{24}$$

$$\mathbf{V}_e = \mathbf{U} + \frac{1}{1 + k_e^2} \left[-\frac{k_e}{B}\mathbf{E'} + \left(\frac{k_e}{B}\right)^2 \mathbf{E'} \times \mathbf{B} - \left(\frac{k_e}{B}\right)^3 (\mathbf{E'}.\mathbf{B})\mathbf{B} \right] \tag{25}$$

where $\mathbf{E'} = \mathbf{E} + \mathbf{U} \times \mathbf{B}$, and $k_i = \frac{\Omega_i}{\nu_{in}}$ and $k_e = \frac{\Omega_e}{\nu_{en}}$ are ion and electron mobility parameters, with $\Omega_i = eB/m_i$ and $\Omega_e = eB/m_e$ being ion and electron gyro frequencies; $\nu_{in} = 2.6 \times 10^{-15}(n_n + n_i)(M_n')^{1/2}$ and $\nu_{en} = 5.4 \times 10^{-16}n_n T_e^{1/2}$ are ion-neutral and electron-neutral collision frequencies, with M_n' denoting mean molecular mass, and densities in m^{-3}.

Altitude variations of collision frequencies (ν_{in} and ν_{en}), gyro frequencies (Ω_i and Ω_e), mobility parameters (k_i and k_e), and conductivities σ_p, $\sigma_\parallel$ and σ_H are shown in Figure 6. As shown, $k_i \ll 1$ at altitudes below about 100 km, and therefore $\mathbf{V_i} \approx \mathbf{U}$ (from equ. 24). Hence ion velocity in the bottom part of E region is determined purely by neutral wind (ions being closely coupled to neutral gas). At altitudes between about 100 and 115 km, k_i is still much less than 1, and so $\mathbf{V}_i \approx \mathbf{U} + k_i\frac{\mathbf{E}}{B}$ where we have used the fact that

15

230 $(k_i/B) \mid \mathbf{U} \times \mathbf{B} \mid \ll \mid \mathbf{U} \mid$. Thus, the relative ion-neutral velocity in the middle part of

231 E region (100-115 km) is parallel to the electric field $\mathbf{E}$.

At high altitudes (above about 180, F region), $k_i \gg 1$, and we can consider the relative

ion-neutral velocity parallel to $\mathbf{B}$ and perpendicular to $\mathbf{B}$. The parallel velocity

$$(\mathbf{V}_i)_\parallel = (\mathbf{U})_\parallel + \left(\frac{k_i}{B}\right) \mathbf{E}_\parallel \tag{26}$$

becomes proportional to the $\mathbf{E}$ field component along $\mathbf{B}$, and perpendicular velocity becomes

$$(\mathbf{V}_i)_\perp = \mathbf{U} + \left(\frac{\mathbf{E}' \times \mathbf{B}}{B^2}\right)_\perp = \frac{\mathbf{E} \times \mathbf{B}}{B^2} \tag{27}$$

For electrons, $k_e \gg 1$ at all altitudes, and therefore

$$(\mathbf{V}_e)_\perp = \frac{\mathbf{E} \times \mathbf{B}}{B^2} \tag{28}$$

232 Hence the ion and electron velocities in the F region (equs 27 and 28) are along the $\mathbf{E} \times \mathbf{B}$

233 direction and independent of neutral wind.

5.2. Ionospheric conductivities

The current density $\mathbf{j}$ at a given height in the ionosphere is given by $\mathbf{j} = ne(\mathbf{V}_i - \mathbf{V}_e)$ in

which singly charged ions and charge neutrality are assumed. Inserting the values of $\mathbf{V}_i$

and $\mathbf{V}_e$, and considering $\mathbf{E} = \mathbf{E}_\parallel + \mathbf{E}_\perp$, the current density becomes

$$\mathbf{j} = \sigma_0 \mathbf{E}_\parallel + \sigma_1 \mathbf{E}_\perp - \sigma_2 \frac{\mathbf{E} \times \mathbf{B}}{B}. \tag{29}$$

$$\sigma_0 = \frac{ne}{B}(k_e + k_i) \tag{30}$$

$$\sigma_1 = \frac{ne}{B}\left(\frac{k_e}{1 + k_e^2} + \frac{k_i}{1 + k_i^2}\right) \tag{31}$$

$$\sigma_2 = \frac{ne}{B}\left(\frac{k_e^2}{1 + k_e^2} - \frac{k_i^2}{1 + k_i^2}\right) \tag{32}$$

234 are the parallel or direct conductivity (σ_0), transverse or Pedersen conductivity (σ_1) and

235 Hall conductivity (σ_2). The parallel conductivity σ_0 corresponds to the conductivity when

16

E is parallel to **B**. When **E** and **B** are perpendicular, the conductivity parallel to **E** (but perpendicular to **B**) is the Pedersen conductivity σ_1, and that perpendicular to both **E** and **B** is Hall conductivity σ_2. If the Hall current is inhibited, then the Pederson current (in the direction of the electric field) will be enhanced due to the polarization effect in the Hall direction. The Cowling conductivity (σ_3) then corresponds to the enhanced current in the direction of the electric field. The narrow current bands of the equatorial electrojet and of the two auroral electrojets owe their existence to the Cowling conductivity acting in these regions.

Spatial variation of conductivity and electrojets

Considering the cartesian coordinate system (x, y, z), with x toward magnetic north, y toward east and z toward Earth's center, the electric field, in general, has components in all directions and magnetic field has components in the x and z directions so that

$$\mathbf{E} = E_x\mathbf{x} + E_y\mathbf{y} + E_z\mathbf{z} \tag{33}$$

$$\mathbf{B} = B\left(\cos I\mathbf{x} + \sin I\mathbf{z}\right) \tag{34}$$

where I is dip angle. The current density has components in all directions and can be simplified to the tensor form

$$\begin{pmatrix} j_x \\ j_y \\ j_z \end{pmatrix} = \begin{pmatrix} \sigma_1 \sin^2 I + \sigma_0 \cos^2 I & -\sigma_2 \sin I & (\sigma_0 - \sigma_1) \sin I \cos I \\ \sigma_2 \sin I & \sigma_1 & -\sigma_2 \cos I \\ (\sigma_0 - \sigma_1) \sin I \cos I & \sigma_2 \cos I & \sigma_1 \cos^2 I + \sigma_0 \sin^2 I \end{pmatrix} \begin{pmatrix} E_x \\ E_y \\ E_z \end{pmatrix} \tag{35}$$

At high latitudes where $I \approx 90°$ and **B** is in the z direction, the above equation reduces to

$$\begin{pmatrix} j_x \\ j_y \\ j_z \end{pmatrix} = \begin{pmatrix} \sigma_1 & -\sigma_2 & 0 \\ \sigma_2 & \sigma_1 & 0 \\ 0 & 0 & \sigma_0 \end{pmatrix} \begin{pmatrix} E_x \\ E_y \\ E_z \end{pmatrix} \tag{36}$$

17

Now suppose the northward current is inhibited ($j_x = 0$), then $E_x = \frac{\sigma_2}{\sigma_1} E_y$ and current in the east-west direction (j_y), which is auroral electrojet current, becomes

$$j_y = \left(\sigma_1 + \frac{\sigma_2^2}{\sigma_0} \right) E_y = \sigma_A E_y. \tag{37}$$

$\sigma_A = \left(\sigma_1 + \frac{\sigma_2^2}{\sigma_0} \right)$ is the auroral electrojet conductivity.

In equatorial region, where I is zero, the current tensor (equ. 35) reduces to

$$\begin{pmatrix} j_x \\ j_y \\ j_z \end{pmatrix} = \begin{pmatrix} \sigma_0 & 0 & 0 \\ 0 & \sigma_1 & -\sigma_2 \\ 0 & \sigma_2 & \sigma_1 \end{pmatrix} \begin{pmatrix} E_x \\ E_y \\ E_z \end{pmatrix} \tag{38}$$

Now if the vertical current j_z is inhibited, then the above equation gives $E_z = -\frac{\sigma_2}{\sigma_1} E_y$ and the east-west current j_y (equatorial electrojet current) becomes

$$j_y = \left(\sigma_1 + \frac{\sigma_2^2}{\sigma_1} \right) E_y = \sigma_3 E_y. \tag{39}$$

$\sigma_3 = \sigma_1 + \frac{\sigma_2^2}{\sigma_1}$ is known as Cowling conductivity. Hence, the auroral electrojet and equatorial electrojet result from the polarization enforced by the restrictions $j_x = 0$ and $j_z = 0$.

6. Ionospheric variations and irregularities

As understood from theory, ionospheric density should vary with time of the day, day of year, season of year and phase of solar cycle. These variations are referred to as diurnal, day-to-day, seasonal and solar cycle variations. In addition, following geomagnetic storms, the ionospheric density increase/decrease very much from its average level, which are known as positive/negative ionospheric storms. The studies of geomagnetic and ionospheric storms are extremely valuable for *space weather* applications. The specific space weather effects include ionospheric density changes that affect GPS applications (communication and navigation), ionospheric currents that affect power supply and gas supply

18

systems, atmospheric heating that affects satellite orbits, and heat and energy transfer that affect system dynamics.

Dense patches of ionization often occurring at E region heights (100-130 km), which is not related to normal E region, is called *sporadic E* or E_s as it occurs sporadically. Sometimes, E_s appears in sheets which hide the overlying F layer; it may also be patchy and partially transparent to radio waves. The horizontal extend of E_s varies from tens to hundreds of kilometers, it drifts with velocities of the order of 50 m s^{-1} and has lifetimes of tens of minutes to several hours. For a theory of E_s see Whitehead (1989).

The most interesting physics in the equatorial F region takes place during postsunset hours when F layer is temporarily raised to high altitudes where collisions are low, E region density is reduced rapidly by chemical recombination, and bottomside F region density gradient is increased by loss of photoionization. In short, the postsunset F layer is in a state of delicate equilibrium; it is like lifting and lowering of a system consisting of a heavy fluid resting on top of a light fluid. Any disturbance cause by background noise, neutral winds, gravity waves, or some other source can disturb the equilibrium and generate instabilities (or irregularities) in the plasma. If conditions (especially high altitude of the ionosphere) are favorable, the irregularities will grow and manifest as *spread F and plasma bubbles* (e.g., Kelley, 1989), which can disturb communication. The need for understanding the characteristics and origin of these irregularities, which cover scale sizes from centimeter to hundreds of kilometers, has increased as demands on communication and navigation have grown through satellite systems such as GPS network.

References

Atreya, S. K., J. H. Waite, T. M. Donahue, A. F. Nagy and J. C. McConnell, Theory, measurements and models of the upper atmosphere and ionosphere of Saturn, In *Saturn*, T. Gehrels and M. S. Mathews (ed), University of Arizona Press, Tucson, AZ, pp.239-280, 1984.

Bailey, G. J., and N. Balan, A low-latitude ionosphere-plasmasphere model, in *STEP Handbook*, edited by R. W. Schunk, pp. 173-206, Utah State Univ. Logan, 1996.

Balan, N., I. S. Batista, M. A. Abdu, J. Macdougall, and G. J. Bailey, Physical mechanism and statistics of occurrence of an additional layer in the equatorial ionosphere, *J. Geophys. Res., 103*, 29169, 1998.

Balan, N., S. V. Thampi, K. Lynn, Y. Otsuka, H. Alleyne, B. G. Fejer and M. A. Abdu, F3 layer during penetration electric field, *J. Geophy. Res.,* doi:10.1029/2008JA013206, 2008.

Brekke, A., Physics of the upper polar atmosphere, John Wiley and sons, New York, 1997.

Kelley, M. C., *The Earth's Ionosphere, plasma physics and electrodynamics*, Academic press, INC, New York, 1989.

Knudsen, W. C., A. J. Kliore and R. C. Whitten, Solar cycle changes in the ionization source of the nightside Venus ionosphere, *J. Geophys. Res., 92,* 13391, 1987.

Mahajan, K. K. and J. Kar, Planetary Ionospheres, *Space Sci. Rev., 47,* 303, 1988.

Ratcliffe, J. A., An introduction to the ionosphere and magnetosphere, Cambridge University Press, 1972.

Rishbeth, H., and O. K. Garriott, Introduction to ionospheric physics, Academic press, New York, 1969.

Viggiano, A. A., and F. Arnold, Ion chemistry and composition of the atmosphere, in

302   Volland (1995), Vol, I, 1, 1995.

303   Whitehead, J. D., Recent work on mid-latitude and equatorial sporadic E, *J. Atmos. terr.*

304   *Phys., 51*, 401, 1989.

305   Yonezawa, T., A new theory of formation of the F2 layer, *J. Radio Res. Labs., 3*, 1, 1956.

Table 1: Ionization and dissociation potentials

Species	Ionization V_p(eV)	Ionization λ(Å)	Dissociation V_p(eV)	Dissociation λ(Å)
N$_2$	15.58	796	9.76	1270
O$_2$	12.08	1026	5.12	2422
O	13.61	911		
N	14.54	853		
NO	9.25	1340	6.51	1905
H	13.59	912		
He	24.58	504		

306 Table 1: Ionization and dissociation potentials (and equivalent wavelengths, according to

307 $V_p = h\nu = h\frac{c}{\lambda}$, with h being Plank's constant and ν and λ being photon frequency and

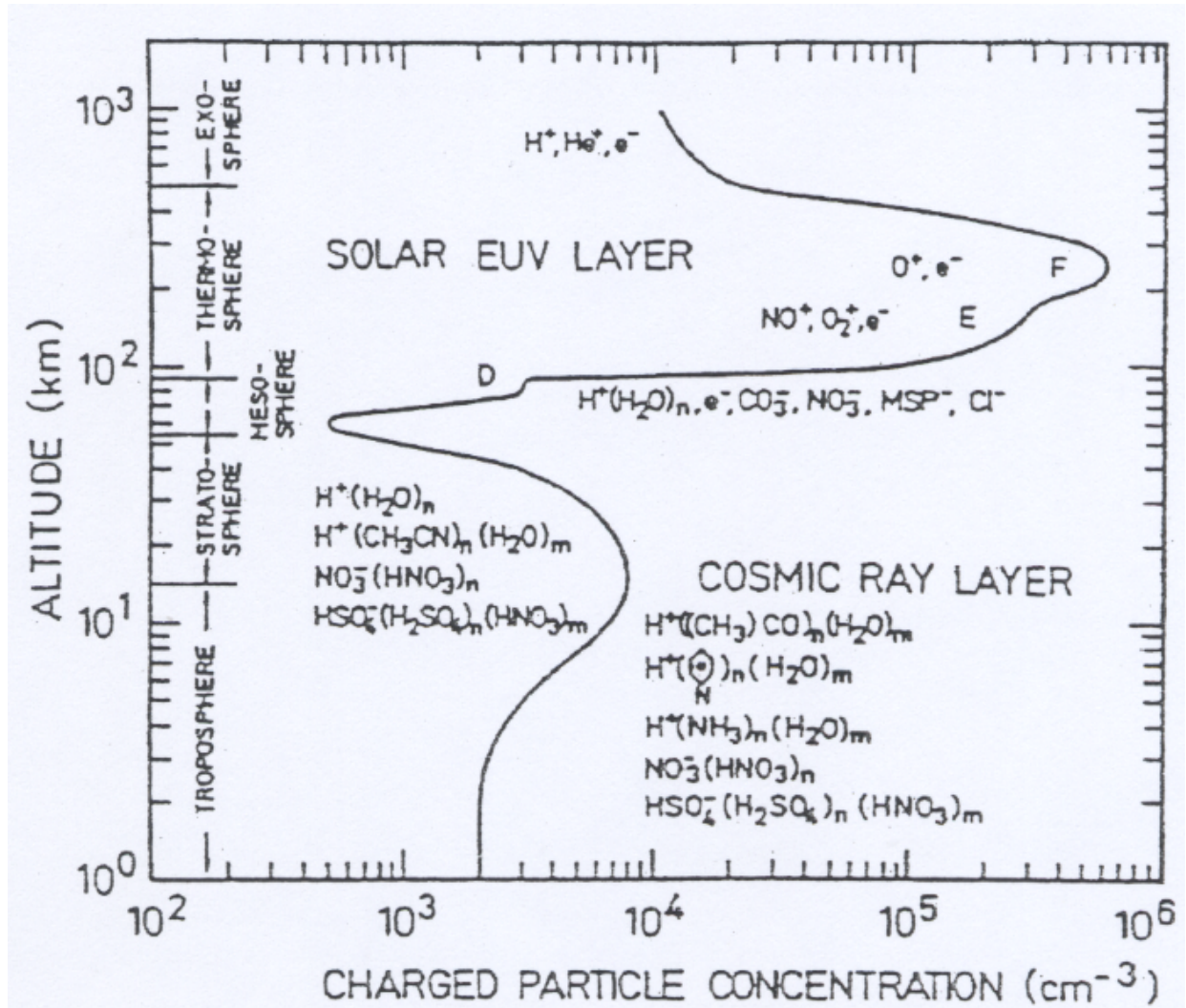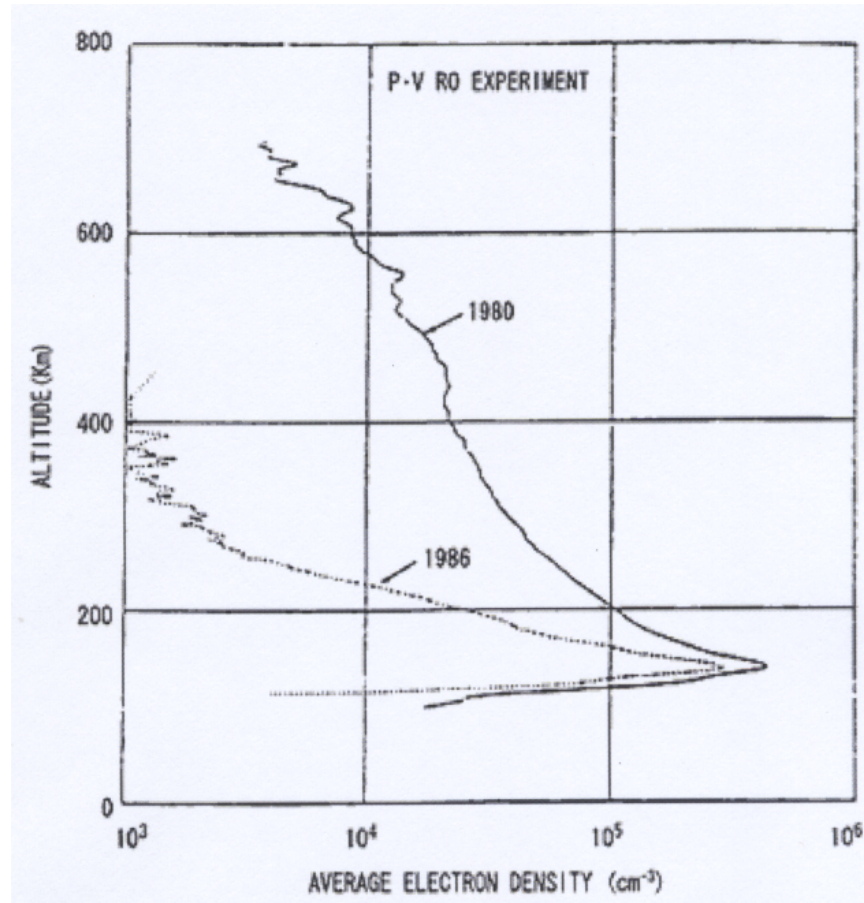308 wavelength) of most abundant species in Earth's atmosphere.

309 Fig. 1. Typical ionization density profile of Earth's atmosphere; major ion species at different altitudes are listed (Viggiano and Arnold, 1995).

311 Fig. 2. Electron density profiles of Venus's ionosphere at solar maximum (1980) and solar minimum (1986) obtained by PVO radio occultation (Knudsen et al., 1987).

313 Fig. 3. Normalized Chapman production function versus reduced height z for different solar zenith angles χ.

315 Fig. 4. The altitude of unit optical depth for wavelengths below 3000Å. The arrows indicate the wavelength regions of typical photoionization and photodissociation.

317 Fig. 5. Ionograms showing the development of F_3 layer at Jicamarca (11.9°S, 76.8°W; dip latitude 1°N) during eastward prompt penetration electric field on 09 November 2004.

319 Fig. 6. Typical altitude profiles of the ion-neutral (ν_{in}) and electron-neutral (ν_{en}) collision frequencies, and ion and electron gyrofrequencies (Ω_i and Ω_e) (a), ion and electron mobility parameters (k_i and k_e) (b), and Pederson, Hall and parallel conductivities (σ_p. σ_H and $\sigma_\parallel$) (c) for an auroral location such as Tromso (69°, 19°E) (Berkke, 1997).

323

Fig. 1: Typical ionization density profile of Earth's atmosphere; major ion species at

different altitudes are listed (Viggiano and Arnold, 1995).

Fig. 2: Electron density profiles of Venus's ionosphere at solar maximum (1980) and solar minimum (1986) obtained by PVO radio occultation (Knudsen et al., 1987).
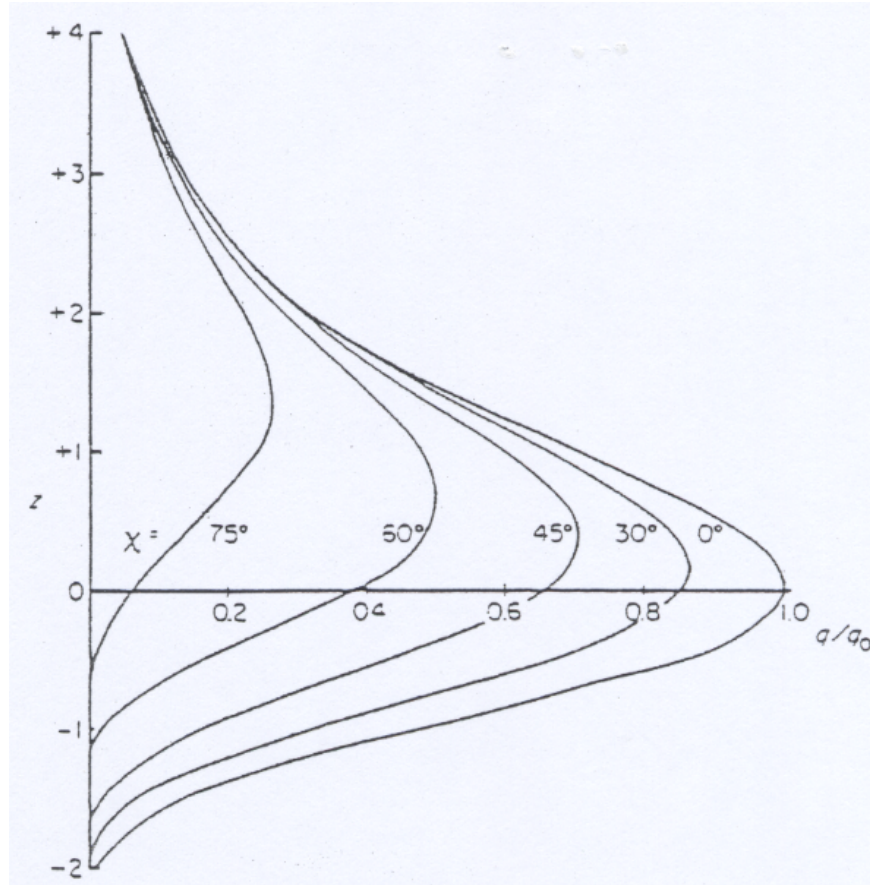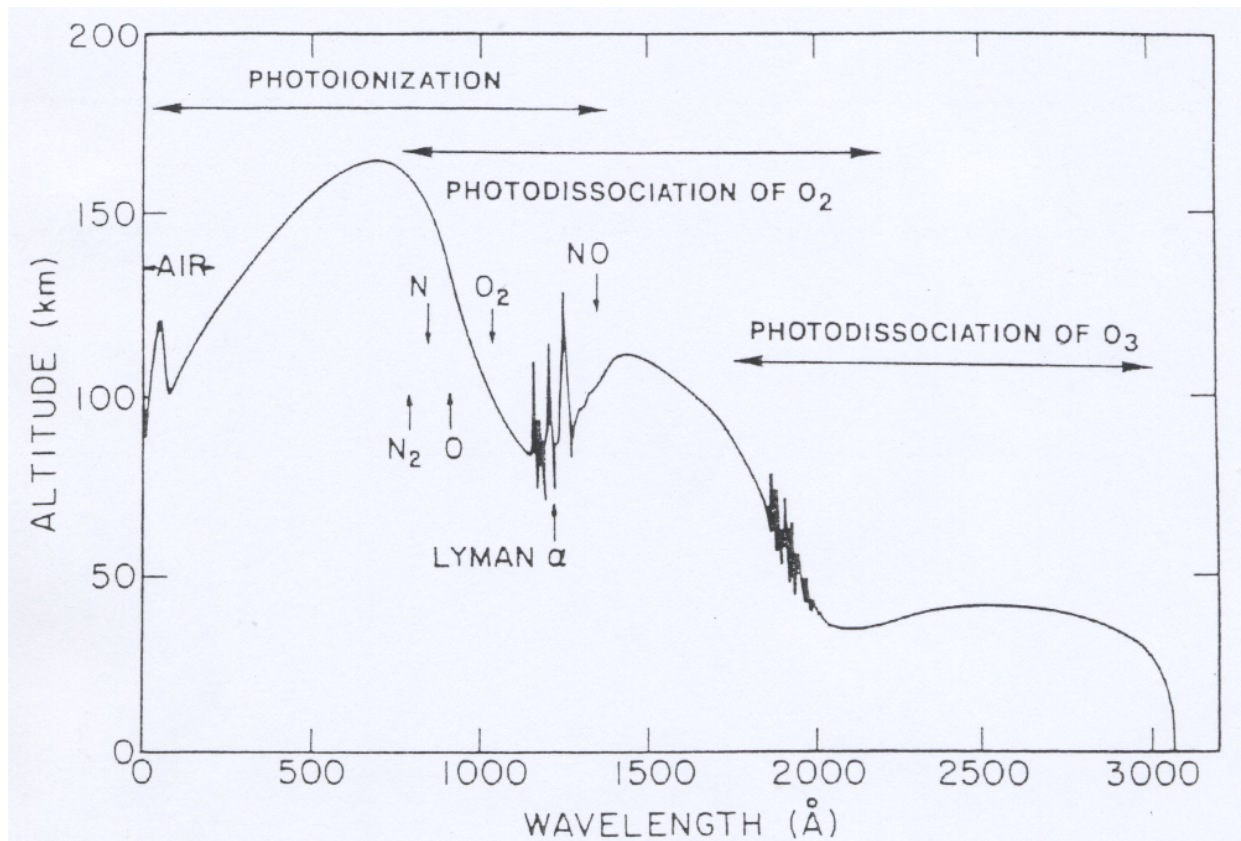
Fig. 3: Normalized Chapman production function versus reduced height z for different

solar zenith angles χ.

Fig. 4: The altitude of unit optical depth for wavelengths below 3000Å. The arrows indicate the wavelength regions of typical photoionization and photodissociation.
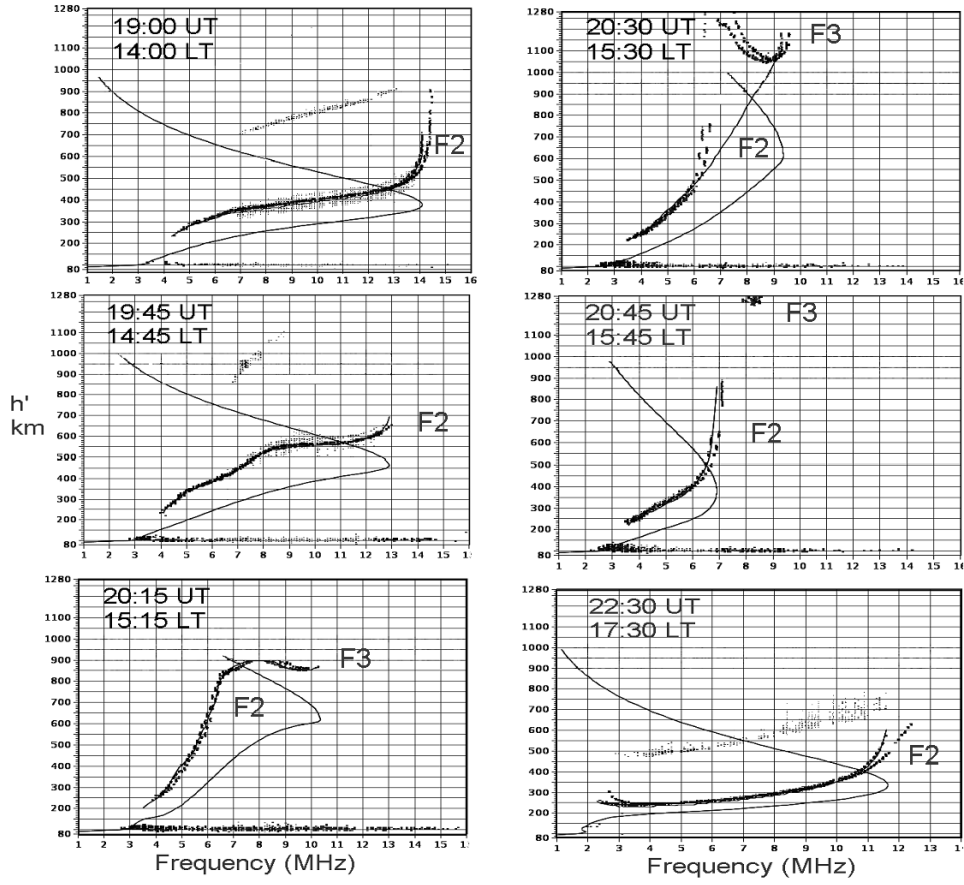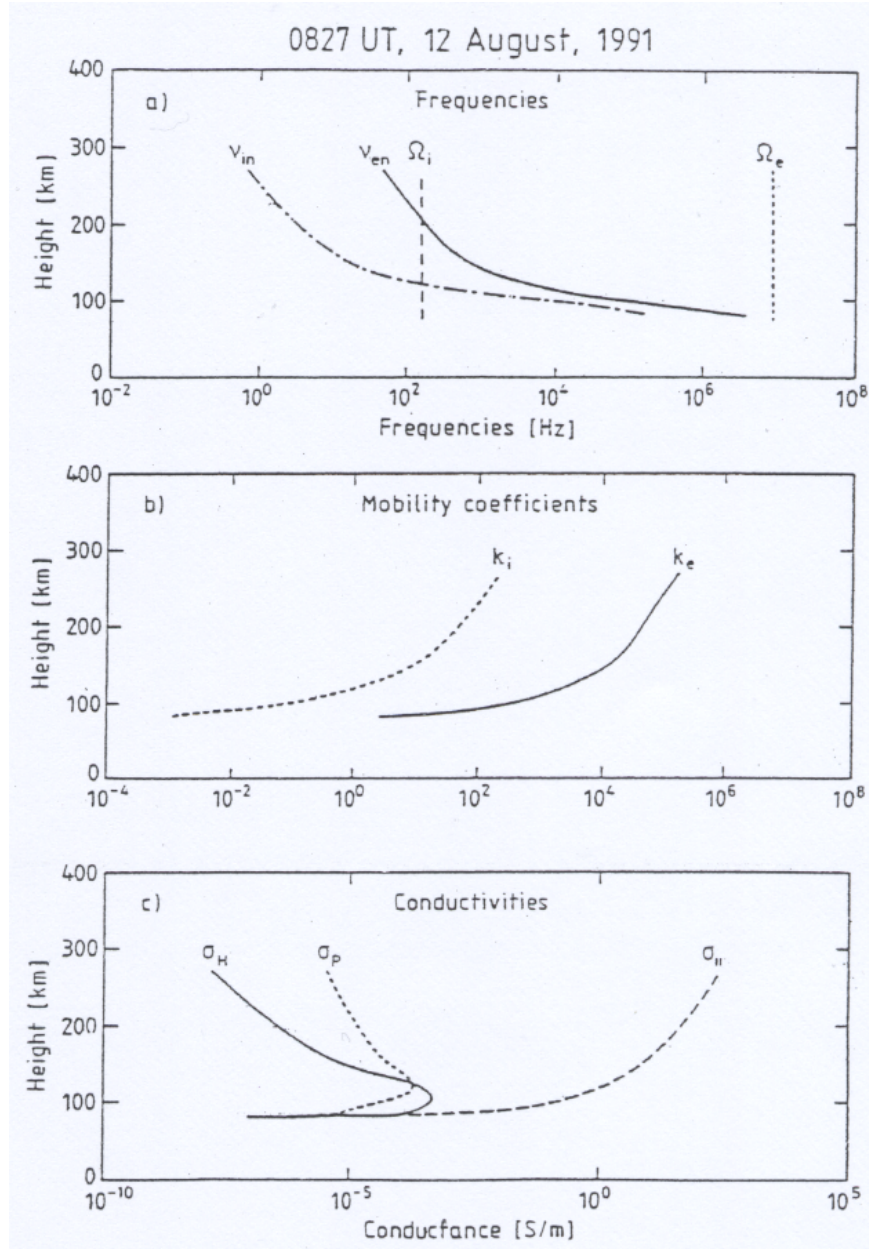
Fig. 5: Ionograms showing the development of F_3 layer at Jicamarca (11.9°S, 76.8°W; dip latitude 1°N) during eastward prompt penetration electric field on 09 November 2004.

Fig. 6: Typical altitude profiles of the ion-neutral (ν_{in}) and electron-neutral (ν_{en}) collision frequencies, and ion and electron gyrofrequencies (Ω_i and Ω_e) (a), ion and electron mobility parameters (k_i and k_e) (b), and Pederson, Hall and parallel conductivities (σ_p, σ_H and $\sigma_\parallel$) (c) for an auroral location such as Tromso (69°, 19°E) (Berkke, 1997).